Empir Econ (2017) 52:255–282
DOI 10.1007/s00181-016-1081-9

The Baltic Dry Index: cyclicalities, forecasting and hedging strategies

Fotis Papailias[1] · **Dimitrios D. Thomakos**[2,3] · **Jiadong Liu**[1]

Received: 1 May 2015 / Accepted: 7 February 2016 / Published online: 6 April 2016

Abstract The cyclical properties of the Baltic Dry Index (BDI) and their implications for forecasting performance are investigated. We find that changes in the BDI can lead to permanent shocks to trade of major exporting economies. In our forecasting exercise, we show that commodities and trigonometric regression can lead to improved predictions and then use our forecasting results to perform an investment exercise and to show how they can be used for improved risk management in the freight sector.

Keywords Baltic Dry Index · Commodities · Concordance · Cyclical analysis · Forecasting · Freights · Hedging · Trade · Turning points

JEL Classification C5 · R00

We are grateful to the Editor, H. Anderson, an Associate Editor and two anonymous referees for valuable comments and suggestions. Any remaining errors are our own.

✉ Fotis Papailias
f.papailias@qub.ac.uk; f.papailias@quantf.com
http://www.quantf.com

Dimitrios D. Thomakos
thomakos@quantf.com; thomakos@uop.gr
http://www.quantf.com

1 Queen's University Management School, Queen's University Belfast, Riddel Hall, 185 Stranmillis, BT9 5EE Belfast, UK

2 Department of Economics, University of Peloponnese, Tripolis, Greece

3 Rimini Centre for Economic Analysis, Rimini, Italy

1 Introduction

The dry bulk shipping sector has long been of interest to investment banks, institutions and academia. The size-specific indices like Supramax, Panamax and Capesize along with the overall Baltic Dry Index (BDI) are being observed daily by economists and market investors. The BDI is defined as a daily weighted average freight price to ship raw materials across the globe in order to be used in the production process. Therefore, it incorporates aspects of the future economic activity and thus has the characteristics of a leading economic indicator. This economic linkage may provide a misleading idea of order regarding the market dynamics of freight rates, as there are at least two typical problems in this market that complicate its dynamics: endogeneity and supply lags. It is straightforward to see that shipping costs affect and are affected by global activity, which means that it is difficult to reliably use even current activity estimates to foretell the path of freight rates. Regarding the supply lags, the supply of freight services (i.e. new and second-hand ships, scrapping etc.) is very inelastic (reacts poorly/slowly to price changes) as it is limited by supply lags and often lacks of market depth, while the demand for freight services tends to be very elastic. Generally, such a set up favours the eventual resolution of supply/demand imbalances and smooths out the path of the BDI.

The transportation of dry bulk goods affects a variety of markets and not just the shipping market. Items such as coal, steel, iron ore, foodstuffs such as corn, wheat and many others indicate that BDI variation should have strong association with the commodities market as well. Recent professional research indicates that this dual causality of BDI and commodities plays a key role in the pricing process. For example, Bornozis (2006) sheds some extra light regarding the global factors that affect the supply and demand in this sector while a report by Giannakoulis and Slorer (2012) reports that the daily crude steel run rate for February 2012 and iron ore imports were surprisingly high. A report by Nomura Equity Research (2012) further mentions excess supply issues with 2012 being the third consecutive year of double-digit supply growth, while demand has never recorded a double-digit growth historically. It is important to notice that these reports expect the BDI to rebound from its current levels.

Later in the paper we also analyse the long-run cumulative effect of BDI to trade. As expected, we find that for major exporters (and importers), such as Australia, Brazil, China, Russia, USA, etc., BDI changes can lead to a permanent shock of various sizes across countries. Therefore, based on our so far discussion, it is understood that BDI is a variable with heavy economic significance. The research purpose of this paper is to: (i) investigate ways and models to accurately forecast this variable and (ii) suggest hedging strategies for those market participants or traders who depend on it.

The academic literature on the BDI and various sub-indices, and the shipping freight rates in general, has a long history, and many papers have analysed various aspects of the behaviour and time series properties of these indices. Driehuis (1970) is among the first to provide a very thorough investigation of the liner freight rates, including a well-formulated economic-theoretical model. Marlow and Gardner (1980) also have an early model on the dry bulk shipping sector, and Beenstock and Vergottis (1989a, b) build an econometric model for the world tanker market and the dry bulk market. In contrast to Adland and Cullinane (2006), Koekebakker et al. (2006) and Batchelor et al.

(2007) are interested in the BDI series as a whole rather than analysing the spot or forward rates separately. More recently, Goulas and Skiadopoulos (2012) analyse the efficiency of the IMAREX futures markets and Lin and Sim (2013, 2014) investigate the relationship and effect of the BDI with trade as well as the transitory negative income shocks impacted by the BDI in Sub-Saharan countries. Furthermore, Lin and Sim (2015) use a BDI-related instrument to estimate the effect of exports on HIV incidence for sub-Saharan countries.

Comparisons of volatility in the dry-cargo ship sector have been conducted by Kavussanos (1996), while the seasonal properties and forecasting in the dry bulk shipping sector are investigated in Cullinane et al. (1999), Kavussanos and Alizadeh (2001) and Kavussanos and Alizadeh (2002).

Using the work of the last mentioned authors as our point of departure, in this paper we are mainly concerned with the cyclical (and not just the seasonal) characteristics of the BDI annual growth series. We test for the cyclical properties of the series, with special emphasis on the frequency part that is closer to the business cycle, and we develop different models to capture and interpret this characteristic. Our analysis shows that there is a strong cyclical pattern of cycle duration of between 3 and 5 years which can be captured by a simple trigonometric regression at these frequencies with relatively good fit. After our cyclical analysis, we consider the problem of forecasting the BDI annual growth series.

Here we take a quite comprehensive stance, compared to the existing BDI-related literature, and we consider a variety of models which incorporate explanatory variables and the cyclical component. Our forecasting exercise is conducted with a focus on the medium- to long-term horizon, and we evaluate 1-, 6- and 12-month-ahead forecasts of the BDI growth. Our results indicate that a considerable proportion of BDI growth variation can be predicted by a combination of explanatory factors and the cyclical pattern that exists in the series. Finally, our forecasting experiments and evaluation further improve the work of Denning et al. (1994).

If the above results are considered as a point for further analysis, a reasonable question to ask next is whether we can use our models for maritime risk management purposes. That is, if cyclicality is indeed present in the data and a model can capture it, then there might be a way of using this, for example, in hedging the path of the BDI or having a portfolio of other assets replicating it or even speculating on its future performance. Investment banks, shipping firms and individual investors that make their business choices based on expectations about the BDI could benefit in terms of correct model timing. We thus go a step further and show how our forecasting results can be put in real-life context in evaluating a straightforward investment strategy: this strategy compares the performance of model-based investment decisions against some alternative benchmarks. The results from this approach indicate that the timing ability of our suggested models works well in an investment-decision context and can thus be further exploited for risk management purposes. The same could be used for freight futures trading.

To sum up, the contribution of this paper to the literature is threefold: (i) analyse the cyclical characteristics of the BDI and its long-run cumulative effect on trade, (ii) exploit these cyclicalities in forecasting and (iii) use the direction of these forecasts to suggest hedging strategies to those exposed to instruments (or prices) related to the BDI.

The rest of the paper is organised as follows. Section 2 presents the data and their descriptive statistics. In Sect. 3, we analyse the long-run cumulative effect of the BDI on trade. Then, in Sect. 4 we present the tests, analysis and discussion on the cyclical behaviour of the BDI annual growth. In Sect. 5, we first introduce the different models we use for predicting the BDI growth, evaluate and discuss their forecasting performance and present the results of the BDI investment strategy which uses the previously described forecasting models. Finally, Sect. 6 offers some concluding remarks.

2 Data

In our analysis, we use monthly data for the BDI and a number of related variables. The full sample range, after adjusting for the computation of annual growth rates, spanned from February 1993 to August 2015 for a total of $T = 271$ monthly observations. The time series is displayed in Fig. 1.

Given the nature of the BDI, it is meaningful to consider commodity variables such as COAL, COPPER, CORN, COTTON, IRON (ore), TIN and WHEAT. In particular, iron ore and coal are the two most important bulk commodities comprising 27 and 26 % of the total dry bulk trade, respectively. Therefore, one would expect the latter variables to be able to predict the BDI.

We also consider CRUDE oil prices (Brent Europe) which might not be transported by bulk shipping but, as we discuss below, has a positive correlation with the BDI (in

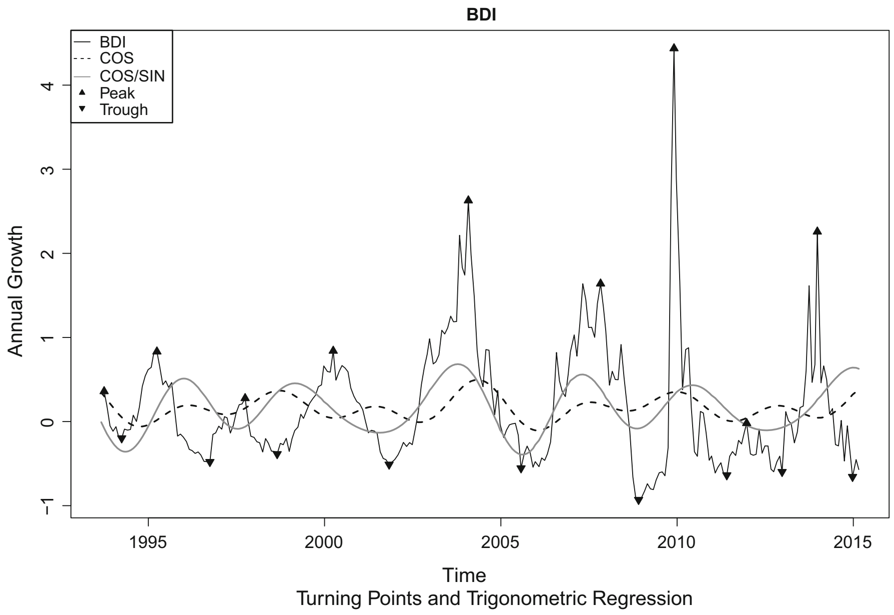

Fig. 1 Annual growth of the BDI with peaks, troughs and the fit of trigonometric regressions

annual growth rates). For the same reason (strong positive of negative correlation), we also include a number of other economic variables such as: the Morgan Stanley global indices for emerging (MSCIEM) and developed (MSCIDEV) markets, the British pound/US dollar exchange rate (GBPUSD), the dollar index (DXY) and, finally, the SPREAD, which denotes the difference between the 10- and 2-year US Treasury yields. All data are collected from Bloomberg, and variables are expressed in annual growth as well.

The basic descriptive statistics of the annual growth series are given in Table 1. In the first panel of the table, we present the statistics for each series, and in the second panel of the table, we have the contemporaneous, full sample, correlations of BDI with the other variables.

The statistics present some interesting features. The BDI has the highest average growth, 16.9 %, among all other variables, except the SPREAD, and this is due to the large increase that it exhibited before the financial crisis in 2008—and also due to the large decrease after the crisis. A similar behaviour is seen in the CRUDE which highlights why CRUDE might be use a potential factor later in the forecasting exercise. It is followed by COPPER and TIN with 9.4 and 9 %, respectively. Of comparable, although smaller, magnitude is the average growth of MSCIEM, TIN, COAL, MSCIDEV, CORN, IRON, WHEAT and COTTON—note that the corresponding standard deviations are almost half that of the BDI for all these variables.

Turning to the correlations, we can see that—in absolute magnitude—the highest correlations are (positive) for the MSCIEM and (negative) for the DXY index, at about 50 %. Note that they all make sense, in that a weaker US dollar was associated with the period of higher global growth, thus higher MSCIEM growth, and the large increase in the BDI. After these variables, we see that TIN, GBPUSD, COPPER and COPPER follow with (positive) correlations above 40 %.

The so far analysis provides some first insights to the relationships of the factors and motivates why those factors are later used in the forecasting experiment.

3 The long-run cumulative effect of BDI to trade

Before we continue our analysis on the cyclical properties and forecasting of the BDI, we provide some extra motivation by highlighting the importance of this index to trade, following the suggestion of a referee.

Table 2 presents the long-run cumulative effect of BDI to total imports and exports series for Australia, Brazil, Canada, China, Euro-Area Aggregate, France, Germany, Italy, Japan, Russia, Spain, UK and US[1].

We follow a very simple yet intuitive methodology by considering the following linear regression model,

$$y_t = \alpha + \beta_1 y_{t-1} + \sum_{j=1}^{s} \gamma_j x_{t-j} + \varepsilon_t \tag{1}$$

[1] The annual growth data for all series spanned from February 1993 to August 2015 (monthly series) and were downloaded using Macrobond Financial

Table 1 Descriptive statistics and correlations of annual growth rates of all variables

	BDI	CRUDE	MSCIDEV	MSCIEM	GBPUSD	DXY	SPREAD	COAL	COPPER	CORN	COTTON	IRON	TIN	WHEAT
Mean	0.169	0.113	0.072	0.090	−0.001	0.007	1.221	0.080	0.094	0.071	0.055	0.061	0.090	0.057
Median	−0.020	0.054	0.113	0.059	0.003	0.007	1.370	−0.038	0.011	0.002	−0.010	−0.090	0.014	0.007
SD	0.707	0.371	0.168	0.287	0.085	0.086	0.918	0.396	0.354	0.352	0.331	0.491	0.321	0.309
Skewness	1.790	0.883	−0.968	0.256	−0.825	0.182	−0.035	1.914	1.188	1.050	1.227	1.435	0.906	1.170
Kurtosis	8.695	4.821	4.208	2.947	4.165	2.440	1.651	7.100	4.895	3.830	5.798	5.363	3.170	4.603
BDI	1													
CRUDE	0.335	1												
MSCIDEV	0.364	0.293	1											
MSCIEM	0.477	0.519	0.666	1										
GBPUSD	0.466	0.322	0.445	0.377	1									
DXY	−0.501	−0.383	−0.184	−0.301	−0.716	1								
SPREAD	0.097	−0.106	−0.184	0.033	0.011	−0.215	1							
COAL	0.264	0.417	0.090	0.226	0.290	−0.431	0.113	1						
COPPER	0.456	0.536	0.440	0.622	0.486	−0.482	−0.012	0.394	1					
CORN	0.022	0.204	0.077	0.204	0.198	−0.334	0.022	0.379	0.165	1				
COTTON	0.332	0.284	0.251	0.345	0.294	−0.458	0.323	0.231	0.420	0.380	1			
IRON	0.336	0.657	0.461	0.650	0.106	−0.412	0.694	0.860	0.702	0.244	0.644	1		
TIN	0.470	0.438	0.416	0.469	0.584	−0.629	0.129	0.694	0.595	0.494	0.543	0.700	1	
WHEAT	0.069	0.350	0.050	0.253	0.285	−0.422	0.026	0.380	0.281	0.714	0.478	0.163	0.488	1

Table 2 Long-run cumulative effect (Ω ratio) of BDI to trade

Shocks of BDI to trade													
	AUSTRALIA	BRAZIL	CANADA	CHINA	EA	FRANCE	GERMANY	ITALY	JAPAN	RUSSIA	SPAIN	UK	USA
s	Imports												
1	0.042	0.176	0.073	0.149	0.082	0.095	0.042	0.077	0.199	0.390	0.045	0.048	0.150
2	0.052	0.199	0.077	0.155	0.074	0.093	0.052	0.092	0.198	0.450	0.060	0.054	0.153
3	0.062	0.215	0.073	0.160	0.096	0.105	0.050	0.095	0.199	0.474	0.066	0.049	0.152
4	0.064	0.238	0.078	0.181	0.066	0.103	0.058	0.102	0.179	0.486	0.066	0.064	0.147
5	0.065	0.233	0.077	0.176	0.050	0.104	0.061	0.103	0.181	0.477	0.079	0.052	0.141
6	0.068	0.228	0.068	0.170	0.033	0.099	0.063	0.101	0.184	0.422	0.071	0.062	0.135
7	0.063	0.233	0.065	0.178	0.028	0.094	0.066	0.103	0.166	0.390	0.071	0.067	0.134
8	0.048	0.239	0.039	0.177	0.045	0.110	0.049	0.091	0.152	0.368	0.058	0.063	0.141
9	0.049	0.245	0.053	0.179	0.050	0.092	0.061	0.097	0.171	0.321	0.066	0.064	0.142
10	0.042	0.242	0.046	0.178	0.020	0.080	0.048	0.086	0.172	0.150	0.056	0.059	0.131
11	0.050	0.235	0.033	0.175	0.017	0.046	0.050	0.070	0.162	0.014	0.052	0.062	0.124
12	0.056	0.256	0.040	0.183	0.035	0.123	0.059	0.100	0.179	0.042	0.062	0.091	0.129
Median	0.054	0.234	0.067	0.176	0.047	0.097	0.055	0.096	0.179	0.390	0.064	0.062	0.141
	Exports												
1	0.117	0.117	0.042	0.135	0.031	0.032	0.039	0.049	0.174	0.540	0.026	0.032	0.107
2	0.123	0.132	0.046	0.132	0.021	0.046	0.047	0.058	0.180	0.557	0.033	0.039	0.113
3	0.130	0.133	0.038	0.146	0.048	0.046	0.047	0.060	0.162	0.588	0.031	0.042	0.113
4	0.145	0.154	0.045	0.147	0.016	0.036	0.054	0.069	0.131	0.536	0.036	0.049	0.114
5	0.160	0.159	0.055	0.146	−0.003	0.044	0.054	0.068	0.118	0.534	0.037	0.053	0.114
6	0.160	0.153	0.046	0.156	−0.019	0.034	0.050	0.068	0.093	0.452	0.030	0.047	0.114

Table 2 continued

Shocks of BDI to trade													
	AUSTRALIA	BRAZIL	CANADA	CHINA	EA	FRANCE	GERMANY	ITALY	JAPAN	RUSSIA	SPAIN	UK	USA
7	0.155	0.174	0.038	0.162	−0.015	0.037	0.058	0.068	0.088	0.344	0.036	0.065	0.109
8	0.146	0.175	0.020	0.149	−0.007	0.037	0.043	0.062	0.075	0.343	0.024	0.049	0.114
9	0.142	0.190	0.036	0.148	−0.001	0.038	0.056	0.070	0.098	0.337	0.040	0.060	0.116
10	0.152	0.191	0.030	0.154	−0.032	0.024	0.051	0.064	0.080	0.341	0.030	0.068	0.115
11	0.154	0.183	0.011	0.150	−0.042	0.024	0.041	0.053	0.063	0.207	0.023	0.070	0.110
12	0.162	0.207	0.022	0.165	−0.016	0.036	0.052	0.067	0.050	0.228	0.032	0.084	0.107
Median	0.149	0.167	0.038	0.148	−0.005	0.037	0.050	0.065	0.095	0.398	0.032	0.051	0.114

for $s = 1, \ldots, 12$, where y_t denotes the imports or exports target variable for a given country and x_t denotes the BDI. Then, we report the ratio,

$$\Omega_s = \sum_{j=1}^{s} \gamma_j/(1 - \beta_1) \tag{2}$$

for each lag s. Notice that the nominator of the Ω_s ratio is the cumulation of the γ_j parameters and, thus, the ratio indicates the long-run cumulative effect of x_t, which is the BDI, on y_t, which is a trade-related series. Consequently, large values of Ω_s across lags denote a larger permanent shock of the BDI to the trade variable.

Table 2 shows that the BDI has a considerable permanent effect for Australia, Brazil, China, Russia and the US. This is expected given that these countries are top exporters in commodities such as Iron Ore, Coal, Corn and Tin, which are shipped in dry bulk. Similarly, the BDI has a considerable permanent effect to total imports for Brazil, China, Japan, Russia and the USA which are major importers of various goods that go into manufacturing production.

This simple exercise illustrates the importance of the BDI for major trading economies and therefore offers additional motivation as to why the BDI must be accurately forecasted in terms of macroeconomic policy and the associated business cycle.

4 Cyclical analysis of the BDI annual growth

4.1 Identification of turning points and tests of synchronicity

In our analysis, we consider the results in Harding and Pagan (2006), where a coherent methodology is presented for testing cycle synchronicity. The testing methodology proposed therein presupposes that one has available indicator variables that identify expansion and contraction periods for each series. There are various ways of getting these indicator variables, but here we follow a straightforward approach as presented in Harding (2008). We briefly summarise the methodology below while full details can be found in the above papers.

Consider a time series of interest y_t and suppose that we would like to find its local turning points (local maxima and minima) in a window of k observations. Then, these local peaks and troughs are given by,

$$\begin{aligned} \wedge_t &\stackrel{\text{def}}{=} I\left[(y_{t-k}, \ldots, y_{t-1}) < y_t > (y_{t+1}, \ldots, y_{t+k})\right] \\ \vee_t &\stackrel{\text{def}}{=} I\left[(y_{t-k}, \ldots, y_{t-1}) > y_t < (y_{t+1}, \ldots, y_{t+k})\right] \end{aligned} \tag{3}$$

where $I(\cdot)$ is the indicator function. While these two variables can be used to mark expansions and contractions, they have the problem that cycle phases may not alternate and, to alleviate this problem, a form of censoring can be used. To do so, one uses the following recursion to construct a single binary variable that marks expansions and contractions and has the cycle phases alternating,

$$S_t \stackrel{\text{def}}{=} S_{t-1}(1 - \wedge_{t-1}) + (1 - S_{t-1}) \vee_{t-1} . \tag{4}$$

Based on the above series, the alternating turning points are then given by,

$$\begin{aligned} \wedge_t^a &\stackrel{\text{def}}{=} S_t(1 - S_{t+1}) \\ \vee_t^a &\stackrel{\text{def}}{=} (1 - S_t)S_{t+1} \end{aligned} \tag{5}$$

The focus of the analysis is then in the S_t series. Consider two such series S_{tx} and S_{ty} for two underlying variables X_t and Y_t; where Y_t denotes the BDI annual percentage change series and X_t denotes another variable which may be commodities, foreign exchange rates and so on. Let $\rho_S \stackrel{\text{def}}{=} \mathsf{Corr}\left[S_{ty}, S_{tx}\right]$ denote the correlation coefficient between the S_{tx} and S_{ty} series. Following Harding and Pagan (2006), the series are said to be in strong positive synchronisation when the following conditions hold,

$$\text{SPS} : \mathsf{E}\left[S_{ty} - S_{tx}\right] = 0 \text{ and } \rho_S \neq 0 \tag{6}$$

where if in addition $\rho_S = 1$ then we have the series to be in strong perfect positive synchronisation. On the other hand, we have that the series are in strong negative synchronisation if they have zero correlation, i.e. when we have,

$$\text{SNS} : \rho_S = 0 \tag{7}$$

without the need to consider the properties of the mean difference $\mathsf{E}\left[S_{ty} - S_{tx}\right]$. Testing the above conditions is easily done via a GMM approach that accounts for the presence of potential heteroscedasticity and autocorrelation. Our results refer to testing these two hypotheses and are all summarised in Table 3. Following Thomakos and Papailias (2014), in addition to the estimates and their z-statistics, we report the (estimate of the) concordance index C which relates to the correlation ρ_S and measures the proportion of time that the two series are in the same phase. Additional details about the structure of the tests and the concordance index can be found in Harding and Pagan (2006). We use two values for k: one corresponding to an annual cycle ($k = 6$ months on either side of the turning point) and one corresponding to a 5-years cycle ($k = 30$ months).

4.2 The turning points of the BDI annual growth

We first look at the visual characteristics of the BDI series and its growth. The peak and trough points are estimated using Eq. (5). The cyclical features of the annual growth of the BDI are evident in Fig. 1.

Notice that for almost a decade (1993–2001), there was an (almost) deterministic cyclical pattern since peaks and troughs occur in similar values for both series and are about equally spaced (this was about the period that the two papers of Kavussanos and Alizadeh (2001) and Kavussanos and Alizadeh (2002) have used in their analysis). From 2001 onwards, it appears that the duration of the cyclical pattern has increased

Table 3 Cyclical analysis and synchronicities of the annual percentage change of the BDI related to other variables

	Mean diff	Mean diff z-stat	Correl	Correl z-stat	C-Index
Annual cycle					
CRUDE	0.050	0.877	0.335	2.727	0.613
MSCIDEV	0.178	1.995	0.364	1.948	0.498
MSCIEM	0.008	0.151	0.477	2.524	0.565
GBPUSD	0.054	0.707	0.466	2.511	0.498
DXY	0.097	1.228	−0.501	−2.339	0.465
SPREAD	0.112	1.425	0.097	0.655	0.509
COAL	0.089	1.359	0.264	2.715	0.649
COPPER	0.035	1.017	0.456	4.674	0.664
CORN	0.008	0.140	0.022	0.150	0.461
COTTON	−0.054	−1.004	0.332	3.104	0.572
IRON	0.158	1.774	0.336	1.618	0.609
TIN	−0.008	−0.135	0.470	3.501	0.616
WHEAT	0.004	0.055	0.069	0.674	0.480
5-years cycle					
CRUDE	0.213	1.899	0.335	1.597	0.435
MSCIDEV	0.223	1.641	0.364	1.952	0.399
MSCIEM	0.374	3.370	0.477	2.953	0.399
GBPUSD	0.038	0.310	0.466	2.407	0.483
DXY	0.095	0.621	−0.501	−3.713	0.262
SPREAD	0.209	1.853	0.097	0.468	0.469
COAL	−0.043	−0.316	0.264	1.456	0.384
COPPER	0.137	1.120	0.456	2.676	0.472
CORN	0.005	0.048	0.022	0.141	0.428
COTTON	−0.213	−2.890	0.332	2.901	0.531
IRON	0.517	5.070	0.116	1.894	0.339
TIN	0.024	0.286	0.470	3.411	0.561
WHEAT	−0.137	−1.654	0.069	0.446	0.428

Entries are the estimates and their z-statistics for the mean differences $\mathsf{E}\left[S_{ty} - S_{tx}\right]$, the correlation ρ_S and the concordance index I. The z-statistics are based on GMM standard errors with automatic lag selection. The null hypothesis for strong positive synchronisation corresponds to $\mathsf{E}\left[S_{ty} - S_{tx}\right] = 0$ and the null hypothesis of strong negative synchronisation corresponds to $\rho_S = 0$. $\mathbf{y}$ denotes the 12-month change of the BDI, and $\mathbf{x}$ denotes each of the variables in the Table

and there is a break in the systematic seasonal behaviour, although as we see later the longer-term cyclical behaviour is still there.

Prying a bit more into the behaviour of the peaks and troughs on the annual growth series, we estimate that, for $k = 6$, the average amplitude during the expansion part of the cycle was about 0.75 % (percentage points) while the average amplitude during the contraction part was about −0.85 %. On the other hand, for the larger—and more relevant cycle—of $k = 30$ there is considerable asymmetry in these amplitudes as they were estimated to be 1.18 and −3.19 %, respectively. These numbers are for the full sample that includes the large fluctuations of 2002 onwards, and indicate the average rise and fall of the BDI growth from the trough to the peak and vice versa and can serve as rough initial guides in our subsequent analysis.

The regularity of rise and fall for the BDI which can be seen by the above identification of the turning points prompts us to consider trigonometric regression models later in the forecasting exercise.

4.3 Coincidence and synchronisation of the BDI annual growth

The presence of a, possibly regular, cyclical component is of interest but lacks further information that can be useful for decision-making and forecasting. To do this, we now turn to the statistics presented previously and see whether the cycles in BDI growth move together with those of other, related, variables.

In Table 3, we have some statistics on the coincidence and possible synchronisation of the annual change of the BDI with the annual change in a number of such variables. Two cycles are considered: an annual and a 5-year cycle, and there are three measures of the degree of synchronisation: the mean difference, the correlation and the C-index (concordance index). Using the same notation as before, Y_t denotes the annual percentage change of the BDI and X_t stands consecutively for each of the following variables: CRUDE, MSCIDEV, MSCIEM, GBPUSD, DXY, SPREAD, COAL, COPPER, CORN, COTTON, IRON, TIN, WHEAT—all defined in the data Section. As mentioned earlier, the C-index is a practically useful measure as it allows us to see which variables have the strongest connection with the BDI.

We can see that the variables that have the highest coincidence with the BDI annual percentage change include the COPPER, COAL, TIN, CRUDE, IRON, COTTON and MSCIEM in the annual cycle and TIN and COTTON in the 5-year cycle—being in phase with the BDI more than 50 % of the time; in particular, COPPER and COAL are in phase more than 65 % in the annual cycle. This is a finding that conforms with intuition, as these variables are commodities whose freight prices have feedback with the BDI itself.

Then, the variables change based on the cycle length we consider. For the annual cycle, we see that MSCIDEV and GBPUSD have a C-index of almost 50 %, while for the 5-year cycle the next important variables are the GBPUSD, COPPER and SPREAD which have a C-index of more than 45 %. Again, these variables conform to the underlying intuition of the factors that affect the BDI: we have the emerging and developed market indices that can be thought as economic strength indicators which move in a procyclical fashion with the BDI and the GBPUSD exchange rate which moves in relative concordance with the BDI.

However, measuring the concordance of the BDI with these other variables—while informative—is not sufficient. We are also interested in the statistical significance of cycle synchronisation. The z-statistics in the tables are for formally testing the hypotheses of positive synchronisation (z-statistic on mean difference) and of negative synchronisation (z-statistic on correlation).

For the annual cycle, and for a 95 % level of significance, we reject the hypothesis of positive synchronisation in favour of a negative one for the MSCIDEV variable.

Then, for CRUDE, MSCIEM, GBPUSD, COAL, COPPER, COTTON and TIN, we reject the negative synchronisation in favour of the positive one. Again, this result is expected due to the relation of the BDI with these variables as previously analysed.

In the 5-year cycle, we see that we reject the hypothesis for positive synchronisation for IRON. This is also expected given the cycle window we are looking in. During the crisis, the BDI was decreasing with IRON not being in the same phase (see the C-Index which is 33.9 % for IRON). On the other hand, and similarly to the annual cycle results, we reject the negative synchronisation hypothesis for GBPUSD, COPPER, COTTON and TIN. For the rest of the variables, we do not have a clear statistical result. It could be argued that from a statistical perspective BDI and variables as MSCIEM, CRUDE, COAL, CORN and WHEAT are asynchronous in the 5-year cycle window.

5 Forecasting the BDI

5.1 Models

The potential presence of cyclicality in the BDI annual growth series, and the presence of variables that are pro- or counter-cyclical with the BDI, both suggest that they might be useful in forecasting the series into the future. Such an exercise goes beyond the relative ability of variables and models to produce (statistically) accurate forecasts and stretches into the realm of practical planning. We thus consider medium- and long(er)-term forecasts that go to 6 and 12 months ahead, horizons that are both practically useful and do not overtax the models that generate the forecasts. In such an exercise, the choice of a benchmark is significant and we could have chosen among a variety of models. However, to ensure that any of the above findings does not bias the final ranking of the models we stick to the standard, a-theoretical, choice of an autoregressive model as the benchmark. We next turn to a presentation and justification of the rest of the models used in our forecasting exercise.

5.1.1 Trigonometric regression

Using the information from the cyclical analysis, we start by considering the simplest model that can capture the cyclical patterns, i.e. a trigonometric regression. This model uses a combination of cosines or sines and cosines at pre-specified frequencies to explain the cyclical trends in the data. We use three frequencies for the fitting that correspond to periods of about 3, 4 and 5 year cycles. Then, two trigonometric models are described as,

$$\text{TRIG\#1} : y_t = \alpha + \sum_{j=1}^{3} \beta_j z_{tj} + \varepsilon_t,\ t = 1, \ldots, T, \tag{8}$$

$$\text{TRIG\#2} : y_t = \alpha + \sum_{j=1}^{3} \left(\beta_j z_{tj} + \gamma_j w_{tj}\right) + \varepsilon_t,\ t = 1, \ldots, T, \tag{9}$$

where $z_{tj} = \cos\left(2\pi\lambda_j t\right)$, $w_{tj} = \sin\left(2\pi\lambda_j t\right)$ are the transcendental factors evaluated at the three chosen frequencies denoted by λ_j. In Fig. 1, we have the full sample fit from the trigonometric model when we fit the first model with cosines and the second model with both sines and cosines. The composite model explains 30 % of the variability of the annual change of the BDI, a rather large number given its simplicity.

These sample fits indicate that the second model is better than the first in capturing the cyclical variability in the BDI annual growth, and hence this is the one we consider for the rest of the study and we denote it by TRIG in what follows.

5.1.2 Factor selection via principal components

Next, we consider some models that are standard choices in the forecasting literature. One of the most widely used class of forecasting methods using variable reduction are factor methods. Factor methods have been at the forefront of developments in forecasting with large data sets and in fact started this literature with the influential work of Stock and Watson (2002a). The defining characteristic of most factor methods is that relatively few summaries of the large data sets are used in the forecasting equation, which thereby becomes a standard forecasting equation as it only involves a few variables. The assumption is that the co-movements across the indicator variables x_t, where $x_t = (x_{t1} \ldots x_{tN})'$ is a vector of dimension $N \times 1$, can be captured by a $r \times 1$ vector of unobserved factors $F_t = (F_{t1} \ldots F_{tr})'$, i.e.

$$\tilde{x}_t = \Lambda' F_t + e_t, \tag{10}$$

where $\tilde{x}_t$ may be equal to x_t or may involve other variables such as, e.g. lags and leads of x_t and Λ is a $r \times N$ matrix of parameters describing how the individual indicator variables relate to each of the r factors, which we denote with the terms 'loadings'. In Eq. (10), e_t denotes a zero-mean $I(0)$ vector of errors that represents for each indicator variable the fraction of dynamics unexplained by F_t, the 'idiosyncratic components'. The number of factors is assumed to be small, meaning $r < \min(N,\ T)$. The main difference between different factor methods relates to how Λ is estimated.

The use of principal component analysis (PCA) for the estimation of factor models is, by far, the most popular factor extraction method. It has been popularised by Stock and Watson (2002a, b), in the context of large data sets, although the idea had been well established in the traditional multivariate statistical literature. The method of principal components is simple. Estimates of Λ and the factors F_t are obtained by solving,

$$V(r) = \min_{\Lambda, F} \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (\tilde{x}_{ti} - \lambda_i' F_t)^2, \tag{11}$$

where λ_i is a $r \times 1$ vector of loadings that represent the N columns of $\Lambda = (\lambda_1 \ldots \lambda_N)$. One, non-unique, solution of Eq. (11) can be found by taking the eigenvectors corresponding to the r largest eigenvalues of the second moment matrix $\tilde{X}'\tilde{X}$, which then are assumed to represent the rows in Λ, and the resulting estimate of Λ provides the forecaster with an estimate of the r factors $\hat{F}_t = \hat{\Lambda}\tilde{x}_t$. To identify the factors up to a rotation, the variables are usually normalised to have zero mean and unit variance prior to the application of principal components; see Stock and Watson (2002a) and Bai (2003).

PC estimation of the factor structure is essentially a static exercise as no lags or leads of x_t are considered. One alternative is dynamic principal components, which, as a method of factor extraction, has been suggested in a series of papers by Forni, Hallin, Lippi and Reichlin [see, e.g. Forni et al. (2000) among others].

We use the above method to extract the principal component using the following indicator variables: CRUDE, MSCIDEV, MSCIEM, GBPUSD, DXY, COAL, COPPER, CORN, COTTON, IRON, TIN and WHEAT.

It is important to notice again here that, given the previous analysis, the cyclical effect is attempted to be captured using the explanatory variables due to their coincidental relationship with the BDI.

5.1.3 Linear regressions

The bulk of the forecasts is generated from linear models using the variables that capture the cyclical properties of the BDI series, as analysed in the previous section, plus combinations of these variables with the extracted PC factors. We thus consider the following generic regression model,

$$y_t = \alpha + \sum_{j=1}^{K} \beta_j x_{t-j} + \sum_{i=1}^{2} \tau_t(\theta_i) + \varepsilon_t, \quad t = 1, \ldots, T, \tag{12}$$

where y_t denotes the annual growth of the BDI series, x_{tj} is the j-th explanatory variable for $j = 1, 2, \ldots, K$ and $\tau_t(\theta_i)$ is a trend component explained below. We use the following sets of explanatory variables:

- Model PCA: in this model $K = r$, where the first r factors from the use of PC of the previous section. We take r to be that number of factors that estimates at least 90 % of the variance of all variables included in the PC analysis.
- Model COM (commodities): in this model $K = 8$ using the following commodities: CRUDE, COAL, COPPER, CORN, COTTON, IRON, TIN and WHEAT.
- Model CRUDE: in this model $K = 1$ and only the variable CRUDE is used.
- Model COAL: in this model $K = 1$ and only the variable COAL is used.
- Model COPPER: in this model $K = 1$ and only the variable COPPER is used.
- Model COTTON: in this model $K = 1$ and only the variable COTTON is used.
- Model IRON: in this model $K = 1$ and only the variable IRON is used.
- Model TIN: in this model $K = 1$ and only the variable TIN is used.
- Model WHEAT: in this model $K = 1$ and only the variable WHEAT is used.
- Model CORN: in this model $K = 1$ and only the variable CORN is used.
- Model IRONCOAL: in this model $K = 2$ and IRON and COAL are used.

As noted at the beginning of this section, our benchmark model is a simple autoregression that is described as,

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t, \quad t = 1, \ldots, T, \tag{13}$$

As above, the cyclical effect is attempted to be captured using the explanatory variables due to their coincidental relationship with the BDI.

A final comment: the inclusion of monthly dummies did not lead to any significant improvement in forecasting performance, and hence cases where these dummies were used are omitted from the presentation. Furthermore, the use of VAR models does not add any forecasting value and, thus, VAR models are omitted.

5.1.4 Forecast generation, averaging and evaluation

We perform a forecasting exercise using the projection method as described in Stock and Watson (2002a). This method, also known as the direct approach, is more robust in the presence of possible model mis-specification. The forecasts for any model m are then given by,

$$\widehat{y}^{f,m}_{t+h} = z'_t \widehat{\beta}^h, \tag{14}$$

where $\widehat{\beta}^h$ is obtained by regressing y_t on the lagged z_{t-h}, h denoting the forecast horizon.[2] z_t is an appropriately dimensioned vector of variables that come from either the trigonometric regression or the linear models described above.

We then specify the (rolling) estimation period R and the evaluation period P so that a summary of a standard pseudo-out-of-sample forecasting algorithm is given as follows.

1. Use the rolling sample of R observations ($R = T - P - h$).
2. With any method described in this section obtain z_{t-h}, with $t = 1, 2, \ldots, R$.
3. Regress y_t on z_{t-h} and obtain $\widehat{\beta}^h$.
4. Calculate the forecasts of $\widehat{y}^{f,m}_{t+h}$ at periods $t = R + 1, R + 2, \ldots, R + h$ using sequentially the values of the explanatory variables $(z_{t-h+1}, z_{t-h+2}, \ldots, z_t)$ a period $t = R$ and the coefficient estimate $\widehat{\beta}^h$.
5. Repeat steps 2 to 4 by rolling the initial sample one period ahead, i.e. by setting $t = 2, 3, \ldots, R + 1$ in step 2 and accordingly in steps 3 and 4.

Due to limited availability of data, the number of rolling estimation periods has been set to $R = \{90, 180\}$.

At the end of this process, we have gathered a total number of P forecast values for any horizon h from any model m. The forecast errors are then calculated as,

$$\widehat{e}^{f,m}_{t+h} = y_{t+h} - \widehat{y}^{f,m}_{t+h} \tag{15}$$

A final step in the forecast generation is an unbiasedness correction that we effect by adjusting the forecasts by the means of the (recursive) forecast errors. This is done so as to (smoothly) 'correct' the forecasts as time elapses. We do this as follows,

[2] Since only lagged values are used as explanatory variables, we do not have to face the endogeneity problem mentioned in the introduction. We could still have an endogeneity problem if the regression error is autocorrelated and of the same order as the delay h used in the explanatory variables. However, our residual diagnostics do not suggest that our residuals suffer from autocorrelation and, therefore, this potential endogeneity source is not present.

$$\tilde{y}^{f,m}_{t+h} = \delta_h \widehat{y}^{f,m}_{t+h} + (1-\delta_h)\frac{1}{t}\sum_{i=R}^{t-1} \widehat{e}^{f,m}_{i+h} \tag{16}$$

where δ_h is the smoothing factor. We use a sequence that, as h increases, progressively gives less weight to the forecast and more to the mean error factor, that is $\delta_h = [0.95 - 0.05(h-1)]$. After these smoothed forecasts are generated, we have the new set of the forecast errors given by,

$$\tilde{e}^{f,m}_{t+h} = y_{t+h} - \tilde{y}^{f,m}_{t+h}. \tag{17}$$

Once these forecasts errors are available then evaluation statistics of interest can be computed. We are particularly interested in the root-mean-squared forecast error (RMSFE) defined as,

$$\text{RMSFE}(h, m) = \frac{1}{P}\sum_{t=R}^{T-h}\left[\left(\tilde{e}^{f,m}_{t+h}\right)^2\right]^{\frac{1}{2}}. \tag{18}$$

We also calculate the Diebold and Mariano (1995) statistic for testing the predictive accuracy of different models. Here we use the two-sided test where the set of hypotheses is as follows:

- $H_0 : E[d_t] = 0$
- $H_A : E[d_t] \neq 0$,

where d_t is the loss differential defined as,

$$d_t = (e^{f,m_1}_{t+h})^2 - (e^{f,m_2}_{t+h})^2, \tag{19}$$

for two competing models m_1 and m_2; $m_1 \neq m_2$ and $t = R, \dots, T-h$. Then, the Diebold and Mariano (1995) test statistic is given by,

$$S = \frac{\overline{d}}{\left(\widehat{LRV}_{\overline{d}}/P\right)^{1/2}}, \tag{20}$$

with,

$$\overline{d} = \frac{1}{P}\sum_{t=R}^{T-h} d_t, \text{LRV}_{\overline{d}} = \gamma_0 + 2\sum_{j=1}^{\infty}\gamma_j, \gamma_j = \text{cov}\left(d_t, d_{t-j}\right), \tag{21}$$

where $\text{LRV}_{\overline{d}}$ is a consistent estimate of the long-run variance $\sqrt{P}\overline{d}$. Under the null of equal predictive accuracy, the statistic is distributed as $S \sim N(0, 1)$.

The sign success ratio (SSR) is defined as the proportion of instances that the direction of the forecasts from each model is the same to the direction of the actual values and is given by,

$$\mathrm{SSR}(h,m)=\frac{1}{P}\sum_{t=R+1}^{T-h} I\left[\mathrm{sgn}(\Delta y_{t+h})=\mathrm{sgn}(\tilde{y}_{t+h}^{f,m}-y_{t+h})\right], \tag{22}$$

where $sgn(\bullet)$ denotes the sign operator and $I(\bullet)$ is an indicator variable which takes the value 1 if the signs are equal and 0 otherwise.

5.2 Forecasting results and discussion

In Table 4, we report the relative (to the AR(1) benchmark) RMSFE, the p value of the Diebold and Mariano (1995) test statistic and the detailed SSR of all models over the respective evaluation periods P. There are two evaluation periods that are dictated by our choice of rolling windows. The rolling window of 90 months allows us an out-of-sample evaluation period from 2001 to 2015 (170 months), while the rolling window of the 180 months allows us an out-of-sample evaluation period from 2009 to 2015 (70 months). The choice of these evaluation periods is obvious, as the second period includes the post-Lehman collapse period that contains the bulk years of the financial crisis. In reading the table, a value greater than one indicates that the benchmark model is better while a value less than one indicates that the corresponding model is better.

Looking at the top panel in Table 4 one thing stands clear: short-horizon forecasts are no better than the benchmark apart from CRUDE, SPREAD and COAL with a relative RSMFE of 0.977, 0.98 and 0.97, respectively (although not statistically significant) and possibly the WHEAT model for its improved SSR (57.96 % against 54.14 % of the benchmark model). However, once we go beyond the 1-month horizon, the results are drastically different and, here, the usefulness of the explanatory variables comes through. Looking at both the 6- and 12-month horizons we see that all the suggested models outperform the benchmark. COAL and IRONCOAL provide a relative RSMFE of 0.758 and 0.779, respectively, for the 6-month horizon, and TRIG returns a relative RSMFE of 0.596 (the smallest across all models) in the 12-month horizon case.

Both results are not surprising: (i) there is a fundamental relationship between BDI, IRON and COAL which explains the good forecasting performance of the model; (ii) in the 12-month horizon where cyclicalities are more present, the use of TRIG (which as mentioned in the previous Section can easily capture those effects) proves to be more effective. It is important to highlight here that TRIG does not depend on any external variables which might be subject to structural change, and thus it does not impose any fundamental assumptions. Therefore, it could be argued that it is more robust as a model choice. It also has the largest SSR equal to 57.32 % against the second best model, which is COPPER, with an SSR of 54.78 %. The respective SSR of the AR(1) benchmark is just 50.96 %.

If we next turn to the results in the bottom panel of Table 4, we see a qualitatively similar overview—supporting the results from the previous table. Looking at the RMSFE results for the 1-month horizon, we still see that CRUDE, SPREAD and COAL along with IRON return a relative RMSFE of 0.976, 0.987, 0.955 and 0.985, respectively. If we use both IRON and COAL then we see that the IRONCOAL model provides a slightly improved forecasting performance with a relative RMSFE of 0.959 (the smallest).

Table 4 Forecasting exercise for the annual percentage change of the BDI. Reporting averages over 170 and 80 evaluation months for 90 and 180 rolling windows, respectively

	RMSFE			DM p value			SSR		
	1-Step	6-Steps	12-Steps	1-Step	6-Steps	12-Steps	1-Step	6-Steps	12-Steps
Rolling window: 90, evaluation periods: 170									
PCA	1.069	0.846	0.650	0.101	0.048	0.002	53.50%	55.41%	51.59%
COM	1.184	0.895	0.656	0.004	0.084	0.000	53.50%	52.87%	50.32%
TRIG	1.610	0.836	0.596	0.000	0.162	0.001	49.68%	52.23%	57.32%
CRUDE	0.977	0.767	0.630	0.568	0.020	0.001	55.41%	52.87%	54.14%
MSCIDEV	1.001	0.791	0.646	0.980	0.021	0.004	54.78%	49.68%	52.87%
MSCIEM	1.000	0.797	0.647	0.988	0.020	0.003	54.14%	49.04%	54.14%
GBPUSD	1.028	0.838	0.670	0.173	0.017	0.002	53.50%	50.96%	55.41%
DXY	1.028	0.828	0.659	0.354	0.017	0.002	52.23%	54.14%	51.59%
SPREAD	0.980	0.808	0.650	0.220	0.004	0.002	55.41%	52.23%	54.14%
COAL	0.970	0.758	0.627	0.572	0.028	0.002	52.87%	49.68%	52.23%
COPPER	1.015	0.840	0.668	0.315	0.017	0.002	50.32%	50.32%	54.78%
CORN	1.007	0.844	0.650	0.663	0.013	0.001	54.78%	49.04%	53.50%
COTTON	1.020	0.870	0.680	0.277	0.040	0.002	52.87%	51.59%	50.96%
IRON	1.051	0.836	0.659	0.182	0.007	0.002	53.50%	52.23%	51.59%
TIN	1.015	0.821	0.662	0.561	0.047	0.002	52.23%	50.96%	50.96%
WHEAT	1.035	0.868	0.691	0.257	0.131	0.003	57.96%	47.77%	49.68%
IRONCOAL	1.009	0.779	0.628	0.840	0.024	0.001	50.32%	49.68%	48.41%
AR(1)	1	1	1	1	1	1	54.14%	48.41%	50.96%
Rolling window: 180, evaluation period 80									
PCA	1.019	0.920	0.854	0.500	0.235	0.140	43.28%	56.72%	49.25%

Table 4 continued

	RMSFE			DM p value			SSR		
	1-Step	6-Steps	12-Steps	1-Step	6-Steps	12-Steps	1-Step	6-Steps	12-Steps
COM	1.016	0.854	0.749	0.707	0.149	0.035	46.27%	49.25%	49.25%
TRIG	1.485	0.779	0.594	0.038	0.276	0.121	52.24%	50.75%	55.22%
CRUDE	0.976	0.836	0.767	0.220	0.138	0.051	53.73%	52.24%	53.73%
MSCIDEV	0.996	0.826	0.714	0.863	0.177	0.127	52.24%	52.24%	52.24%
MSCIEM	1.006	0.879	0.838	0.653	0.173	0.090	56.72%	52.24%	53.73%
GBPUSD	1.013	0.896	0.872	0.523	0.164	0.152	52.24%	47.76%	52.24%
DXY	1.004	0.899	0.906	0.782	0.165	0.322	49.25%	46.27%	50.75%
SPREAD	0.987	0.865	0.812	0.529	0.138	0.036	50.75%	49.25%	50.75%
COAL	0.955	0.793	0.683	0.394	0.153	0.092	47.76%	53.73%	53.73%
COPPER	1.000	0.878	0.844	0.974	0.147	0.068	49.25%	49.25%	55.22%
CORN	1.005	0.902	0.887	0.716	0.174	0.167	44.78%	46.27%	50.75%
COTTON	0.990	0.887	0.859	0.391	0.157	0.079	50.75%	47.76%	49.25%
IRON	0.985	0.869	0.785	0.297	0.129	0.047	50.75%	50.75%	50.75%
TIN	1.005	0.852	0.780	0.798	0.192	0.102	47.76%	49.25%	46.27%
WHEAT	0.993	0.884	0.898	0.518	0.172	0.283	55.22%	49.25%	49.25%
IRONCOAL	0.959	0.790	0.692	0.448	0.159	0.090	43.28%	49.25%	52.24%
AR(1)	1	1	1	1	1	1	50.75%	47.76%	53.73%

RMSFE denotes the relative root-mean-squared forecast error of each method to the benchmark. DM denotes the p value of the two-sided Diebold and Mariano (1995) statistic using the squared difference of the forecast error of each method relative to the benchmark. SSR denotes the sign success ratio of each method. PCA denotes the principal components method, COM denotes the linear regression model using CRUDE, COAL, COPPER, CORN, COTTON, IRON, TIN and WHEAT explanatory variables. The benchmark model is the AR(1)

TRIG again is the best performer across the 6-month and 12-month horizons which corroborates the earlier analysis on the cyclicality of BDI growth and suggests a continuation of the cyclical path of the earlier sample. TRIG relative RSMFE is equal to 0.779 and 0.594 for the 6-month and 12-month forecasting horizon, respectively. As before, it also returns the largest SSR in the 12-month horizon which is equal to 55.22 % compared to 53.73 % of the benchmark.

Summarising the above discussion, we see that all the suggested variables have potential benefits in the BDI forecasting. If we have to choose a model which depends on explanatory factors to forecast the BDI for mid- to long-term forecasting, we could select IRONCOAL. However, for a more robust forecasting, and based on the analysis above more accurate predictions, it would be meaningful to consider TRIG model.

5.3 Model-based investing and risk management

As noted in the introduction, there is a clear need for adding another step in our analysis: even if we accept the presence of cyclicality in the data and, more so, 'believe' our forecasting models, one needs to see them put into a decision-making context. In this section, we attempt to do that by considering how these models can be put into use for investing in the BDI and/or performing risk management by utilising our forecast track record. If our approach in this section is successful, then it opens up a practical use of the forecasting models and many other ideas can possibly be put to good use for anyone that is interested in the BDI path.

The idea here is very simple: if the signs of the forecasts are accurate (i.e. more accurate than a random sign choice), then we can invest or hedge the BDI by placing an appropriate 'bet', going either long or long/short depending on our risk preferences. If the model sign suggests a rise in BDI's annual return, then we should be 'buying' the BDI, and if the model sign suggests a fall in BDI's return, then we should be 'selling' the BDI or, at least, avoiding exposure in the market. Alternatively, one can hedge the BDI by going (appropriately) long or short in any kind of asset that moves along with the BDI: for example, if the model sign suggests a rise in the BDI's return and we want to cover (hedge) ourselves from a possible mistake, then we should 'buy' the BDI and sell an asset that is positively correlated with the BDI (or buy an asset that is negative correlated with the BDI—the result is qualitatively the same).

Although the BDI is not directly tradable, there are many ways in which one can track its path via tradable assets. For example, one can form a portfolio based on assets that are highly correlated with the BDI or consider future contracts. For illustrating the usefulness of the timing ability of the forecasting models, we proceed as if the BDI was directly investable.

We next describe in some detail the way we conduct our investing experiment. We have, as noted before, two strategies: (i) a 'long only' (L) and (ii) a 'long/short' (LS). Both strategies are evaluated in the following manner:

1. We use a 90-month rolling window as our in-sample period and compute the 12-month-ahead forecast for each of our forecasting models.
2. If the sign of the forecast is positive we open, a new long 'position' on the BDI which we hold for the next 12 months; if the sign is negative, we either stay out of

the market (L) or open a new short ‘position’ on the BDI which we also hold for the next 12 months (LS).

3. We allow the window to roll 1 month ahead, and we repeat the whole procedure; this implies that we are opening one new position each month which stays active for the next 12 months.

The performance of the positions thus obtained is to be compared with three benchmarks: one is the performance based on the forecasted signs of the AR(1) model (which is also our benchmark in the forecasting exercise described in the previous section), the other is the performance of just holding the BDI (and to mimic the above timing procedure we assume that we open a new long ‘position’ for the BDI every month), and the final is the time series momentum of Moskowitz et al. (2012). This last approach is particularly relevant as a benchmark, since it applies a sign-based methodology and it is implemented in a similar fashion with the proposed approach that we take. In particular, in the time series momentum one looks at the past sign of a series of returns of an asset and goes long or long/short based on it. There is, therefore, a similarity but also a crucial difference between the momentum approach and ours: in the former, the past is used and is believed that its sign is propagated into the future, while in the latter, a model-based sign forecast is used.

If there is cyclicality that is being captured by the forecasted signs, then our suggested procedure should be able to illustrate it: when the BDI is forecasted to fall over the next 12 months and we either eliminate our exposure into it or even go against it, then we should do better by just holding on to it. Furthermore, since we are using a long-term forecast that goes into the next year, we should have the trigonometric models perform better than other ones, including all three benchmarks. Our results are given in Table 5, and the cumulative return performance is illustrated in Fig. 2. In the table, we present various statistics on the investment performance of the suggested approach, the difference between the two is the evaluation period—in the second of these tables, the evaluation starts one year before the last financial crisis.

The main result that can be seen immediately from both panels of Table 5 is that there is economic value in the use of model sign-based timing. In particular, the predicted signs that are based on the PCA, COM, TRIG, COAL, IRON and IRONCOAL models have the best economic performance: they have the highest cumulative returns, the highest Sharpe ratios and the low maximum drawdowns.[3]

Looking at the top panel of Table 5, we see that the sign-based performance of the COAL, IRONCOAL and PCA is by far the best. For the long-only approach, the annualised Sharpe ratio exceeds 1.5, which compares to a 1.05 value for the BDI, a 0.369 for the AR(1) model and a 0.428 for TSM. They thus outperform the three benchmarks by a wide margin in terms of risk-adjusted returns. The results become even better when we consider the long-short approach, clearly indicating that there are indeed alternating signs in the future path of the BDI returns which can be exploited via the forecasting models. In particular, we see that COAL and PCA provide a cumulative return of 82.5 and 82.8 % in the long-only approach 120.9 and 121.6 %, respectively, in the long-short approach. These results strongly suggest not only that the forecasting

[3] Values for average, volatility and Sharpe ratio are annualised.

Table 5 Investing in the BDI

	Long					Long-short				
	Average	Volatility	Sharpe	Cumulative	Drawdown	Average	Volatility	Sharpe	Cumulative	Drawdown
Rolling window: 90, evaluation periods: 168 (2001–2015)										
PCA	0.046	0.025	1.838	0.828	0.040	0.061	0.026	2.350	1.216	0.026
COM	0.045	0.025	1.801	0.806	0.039	0.059	0.026	2.251	1.164	0.026
TRIG	0.042	0.026	1.627	0.735	0.063	0.053	0.027	1.939	0.997	0.038
CRUDE	0.044	0.025	1.724	0.772	0.022	0.056	0.027	2.098	1.083	0.026
MSCIDEV	0.042	0.025	1.653	0.728	0.033	0.052	0.027	1.907	0.979	0.030
MSCIEM	0.041	0.025	1.611	0.702	0.023	0.050	0.028	1.801	0.921	0.030
GBPUSD	0.040	0.025	1.600	0.694	0.023	0.049	0.028	1.767	0.902	0.030
DXY	0.040	0.025	1.584	0.691	0.030	0.049	0.028	1.754	0.895	0.027
spread	0.041	0.026	1.594	0.708	0.033	0.051	0.028	1.824	0.934	0.027
COAL	0.046	0.025	1.819	0.825	0.022	0.061	0.026	2.338	1.209	0.026
COPPER	0.042	0.026	1.627	0.719	0.022	0.052	0.028	1.872	0.960	0.027
CORN	0.042	0.025	1.664	0.737	0.027	0.053	0.027	1.948	1.002	0.030
COTTON	0.041	0.026	1.603	0.716	0.049	0.051	0.028	1.860	0.953	0.027
IRON	0.042	0.026	1.648	0.731	0.025	0.053	0.027	1.919	0.986	0.035
TIN	0.041	0.026	1.619	0.714	0.033	0.051	0.028	1.851	0.948	0.030
WHEAT	0.038	0.027	1.410	0.630	0.072	0.044	0.029	1.517	0.763	0.077
IRONCOAL	0.045	0.026	1.773	0.806	0.030	0.059	0.026	2.249	1.163	0.026
AR(1)	0.009	0.023	0.369	0.116	0.182	−0.014	0.031	−0.459	−0.177	0.283
TSM	0.009	0.021	0.428	0.120	0.154	−0.014	0.031	−0.446	−0.172	0.261
BDI	0.032	0.030	1.050	0.504	0.108	0.032	0.030	1.050	0.504	0.108
Rolling window: 90, evaluation period:s 90 (2007–2015)										
PCA	0.032	0.025	1.270	0.315	0.040	0.047	0.028	1.634	0.490	0.028

Table 5 continued

	Long					Long-short				
	Average	Volatility	Sharpe	Cumulative	Drawdown	Average	Volatility	Sharpe	Cumulative	Drawdown
Rolling window: 90, evaluation periods: 168 (2001–2015)										
COM	0.036	0.026	1.396	0.361	0.039	0.054	0.027	1.998	0.597	0.026
TRIG	0.031	0.027	1.174	0.309	0.063	0.046	0.029	1.594	0.478	0.038
CRUDE	0.030	0.025	1.213	0.295	0.022	0.043	0.029	1.487	0.446	0.045
MSCIDEV	0.027	0.025	1.072	0.255	0.033	0.036	0.030	1.200	0.357	0.045
MSCIEM	0.026	0.025	1.077	0.254	0.023	0.036	0.030	1.194	0.355	0.045
GBPUSD	0.025	0.025	1.030	0.241	0.023	0.033	0.030	1.106	0.328	0.045
DXY	0.025	0.025	1.023	0.242	0.030	0.033	0.030	1.112	0.330	0.045
spread	0.026	0.025	1.013	0.247	0.033	0.034	0.030	1.143	0.339	0.045
COAL	0.031	0.025	1.242	0.305	0.022	0.045	0.029	1.560	0.468	0.045
COPPER	0.028	0.025	1.108	0.267	0.022	0.038	0.030	1.285	0.384	0.045
CORN	0.033	0.026	1.267	0.322	0.027	0.048	0.028	1.688	0.507	0.030
COTTON	0.025	0.026	0.970	0.236	0.049	0.032	0.030	1.068	0.316	0.045
IRON	0.034	0.026	1.331	0.342	0.025	0.051	0.028	1.850	0.554	0.026
TIN	0.026	0.025	1.051	0.250	0.033	0.035	0.030	1.163	0.345	0.045
WHEAT	0.024	0.027	0.874	0.224	0.072	0.030	0.030	0.991	0.292	0.077
IRONCOAL	0.035	0.026	1.353	0.352	0.030	0.053	0.028	1.925	0.576	0.026
AR(1)	−0.003	0.021	−0.136	−0.027	0.182	−0.023	0.031	−0.752	−0.185	0.283
TSM	−0.005	0.017	−0.289	−0.043	0.154	−0.027	0.031	−0.891	−0.213	0.261
BDI	0.017	0.031	0.557	0.157	0.108	0.017	0.031	0.557	0.157	0.108

Reporting annualised statistics over 168 and 90 evaluation months. Average denotes the annualised mean of returns. Volatility denotes the annualised standard deviation of returns. Sharpe ratio denotes the annualised average/standard deviation ratio of returns. Cumulative denotes the cumulative return. Drawdown denotes the maximum drawdown of the cumulative return. TSM denotes the time series momentum strategy. BDI denotes the actual percentage change of the BDI series

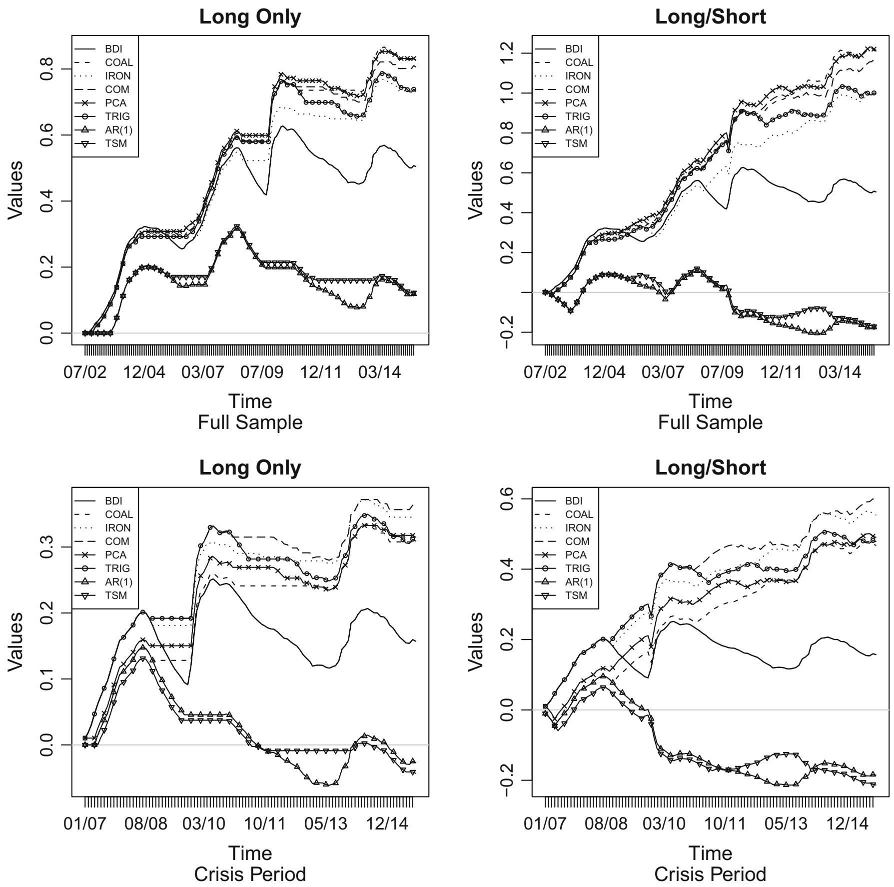

Fig. 2 Investing in the BDI. Comparing the cumulative performance of different strategies using 168 (full sample) and 90 (crisis period) evaluation months

exercise of the previous section was not futile but, on the contrary, it provides us with tools to exploit opportunities in a BDI risk management context.

To further pursue the potential of what is presented above, consider Fig. 2 and notice how, with the onset of the financial crisis in 2008, the return of holding into BDI was falling until 2009. However, the suggested models correctly capture this effect keeping the investor out of the market or indicating a reverse position (long to short).

What if we perform our analysis starting a year before the crisis in 2007? Would we still have been able to reduce our exposure? Although the answer is affirmative, we repeat the analysis and evaluation statistics and present the results in the bottom panel of Table 5 and in Fig. 2. As can be seen clearly from the figure, again the steep fall of the BDI during the crisis is avoided.

Also, the BDI was falling from 2010 to 2013. Even during this time, the suggested models were able to avoid losses and even generate profits (in the long-short approach).

All in all, the results of this section support our earlier findings and, moreover, they transform them into practical tools that can be used from anyone who wishes to manage exposure to the future path of the BDI. There are caveats, obviously, to what we just presented (such as that the BDI is not directly tradable), but the overall good performance of the strategies which are based on model forecasts is such that leaves room for many different ways for further improvements.

6 Conclusions

Our overall analysis provides several interesting and novel results about the evolution of BDI annual growth. First, the contribution of the paper to the literature is the cyclical analysis of the series at different levels. Past research was limited to seasonal analysis (see Kavussanos and Alizadeh (2001, 2002)). We find that there is a strong cyclical pattern of cycle duration of between 3 and 5 years and that this pattern is relatively stable across time.

Second, we perform a comprehensive forecasting performance evaluation by considering a variety of models and model averages that incorporate explanatory variables—carefully selected by top-down elimination—and the cyclical component found in the first part of our analysis. The results of our forecasting exercise show that performance gains are possible when using auxiliary information, either in the form of explanatory variables or in the form of the cyclical component of the BDI. These gains are not uniform across all models examined and are concentrated mainly in the medium- and longer-term forecasting horizons. However, in the cases where outperformance of the benchmark is found we can see several occasions that this is of a rather large magnitude. A judicious choice of models, that incorporate the features that affect the BDI, can thus lead to good forecasting performance and aid in planning and management in using the future direction of the BDI.

Finding cyclicality in economic time series might be considered old-fashioned, but in the present case we cannot refute it easily. Not only the models we present exhibit very good statistical forecasting performance, they can also be used for controlling financial exposure and risk to the BDI. In the last part of our analysis, we perform a risk management experiment where 12-month-ahead forecasts are used to decide whether or not to invest in the BDI. Within the limitations we discussed above, the results on this third part of our analysis strongly support the long(er)-term potential benefits of using sign-based timing for investing or hedging the BDI. Not only do we find that the trigonometric model gives the best economic performance in this experiment, we also find that all of our forecasting models provide a better decision-making tool than any of the three benchmarks we employ.

Our results now open up a very interesting avenue of future research: How can we construct a realisable risk management system that will utilise the model signs and associated information? so that one can exercise a higher degree of control when exposed to BDI fluctuations. Such a system will be used for both investing in and hedging BDI risk and should depend on assets that are immediately available for trading. We are currently pursuing this line of research.

References

Adland R, Cullinane K (2006) The non-linear dynamics of spot freight rates in tanker markets. Transp Res Part E 42:211–224

Bai J (2003) Inferential theory for factor models of large dimension. Econometrica 71:135–173

Batchelor R, Alizadeh AH, Visvikis I (2007) Forecasting sport and forward prices in the international freight market. Int J Forecast 23:101–114

Beenstock M, Vergottis A (1989a) An econometric model of the world shipping market for dry cargo, freight and shipping. Appl Econ 21:339–356

Beenstock M, Vergottis A (1989b) An econometric model of the world tanker market. J Transp Econ Policy 23(3):236–280

Bornozis N (2006) Dry bulk shipping: the engine of global trade. Barron's, New York

Cullinane KPB, Mason KJ, Cape MB (1999) Forecasting the baltic freight index: Box–Jenkins revisited. In: Meersman H, Van de Voorde E, Winkelmans W (eds) World transport research, selected proceedings of the 8th world conference on transport research, vol I., Transport modes and systems. Elsevier Science, Amsterdam, pp 73–86

Denning KC, Riley WB, DeLooze JP (1994) Baltic freight futures: Random walk or seasonally predictable? Int Rev Econ Finance 3(4):399–428

Diebold FX, Mariano RS (1995) Comparing predictive accuracy. J Bus Econ Stat 13:253–263

Driehuis W (1970) An econometric analysis of liner freight rates. Weltwirtschaftliches Archiv 104:96–119

Forni M, Hallin M, Lippi M, Reichlin L (2000) The generalised factor model: identification and estimation. Rev Econ Stat 82:540–554

Giannakoulis F, Slorer O (2012) Global shipping: time to buy containerships?. Morgan Stanley Research, North America

Goulas L, Skiadopoulos G (2012) Are freight futures markets efficient? Evidence from IMAREX. Int J Forecast 28(3):644–659

Harding D (2008) Detecting and forecasting business cycle turning points. In: MPRA Paper No. 33583

Harding D, Pagan A (2006) Synchronization of cycles. J Econom 132:59–79

Kavussanos MG (1996) Comparisons of volatility in the dry-cargo ship sector: spot versus time charters. J Transp Econ Policy 30(1):67–82

Kavussanos MG, Alizadeh AH (2001) Seasonality patterns in dry bulk shipping spot and time charter freight rates. Transp Res Part E 37:443–467

Kavussanos MG, Alizadeh AH (2002) Seasonality patterns in tanker spot freight rates. Econ Model 19: 747–782

Koekebakker S, Adland R, Sodal S (2006) Are spot freight rates stationary? J Transp Econ Policy 40(3): 449–472

Lin F, Sim NCS (2013) Trade, income and the Baltic Dry Index. Eur Econ Rev 59:1–18

Lin F, Sim NCS (2014) Baltic Dry Index and the democratic window of opportunity. J Comp Econ 42: 143–159

Lin F, Sim NCS (2015) Exports, HIV incidence and the Baltic Dry Index: further evidence from sub-Saharan Africa. Econ Lett 126:35–39

Marlow PB, Gardner B (1980) Thoughts on the dry bulk shipping sector. J Ind Econ 29(1):71–84

Moskowitz TJ, Ooi YH, Pedersen LH (2012) Time series momentum. J Finan Econ 104:228–250

Nomura Equity Research (2012) Anchor. report. 02–2012:1–44

Stock J, Watson M (2002a) Forecasting using principal components from a large number of predictors. J Am Stat Assoc 297:1167–1179
Stock J, Watson M (2002b) Macroeconomic forecasting using diffusion indexes. J Bus Econ Stat 20:147–162
Thomakos DD, Papailias F (2014) “Out of Sync”: the breakdown of economic sentiment cycles in the EU. Rev Int Econ 22(1):131–150